Australian Bureau of Statistics

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016

Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

Note: While the 2011 and 2016 Censuses were predominantly the same, there were some minor differences. For example, a number of changes were made to how industry of employment information was collected for the 2016 Census. The ABS advises this data is not directly comparable to 2011 industry data and should not be used to measure longitudinal transitions. For further information refer to Industry of Employment (INDP) in Census of Population and Housing: Understanding the Census and Census Data, Australia, 2016 (cat. no. 2900.0).

Users are also encouraged to read Understanding the data pages in Census of Population and Housing: Understanding the Census and Census Data, Australia, 2016 (cat. no. 2900.0) for information to assist with using and interpreting specific data items across time.

INTRODUCTION

The Census of Population and Housing is conducted every five years to measure the number of people and dwellings in Australia on Census Night. The Census also provides information on the key characteristics of people and dwellings for small geographic areas and small population groups.

The Australian Census Longitudinal Dataset (ACLD) uses data from the Census of Population and Housing to build a rich longitudinal picture of Australian society. The ACLD can uncover new insights into the dynamics and transitions that drive social and economic change over time, and how these vary for diverse population groups and geographies.

The ACLD is a random 5% sample of the Australian population and three waves of data have so far contributed to the ACLD from the 2006 Census (Wave 1), 2011 Census (Wave 2) and 2016 Census (Wave 3).

In this release of the 2011-2016 ACLD, a representative sample of over 1.2 million records from the 2011 Census (Wave 2) was brought together with corresponding records from the 2016 Census (Wave 3) to form the 2011 Panel of the ACLD. The 2011 Panel includes new births and migrants since the 2006 Census, and is a rich source for exploring how Australian society has changed between the 2011 and 2016 Censuses.

A second release of the 2016 ACLD in mid-2018 will include additional data items on the 2011 Panel, as well as an updated 2006 Panel consisting of a linked sample between the 2006, 2011 and 2016 Censuses. The 2006 Panel was first released in December 2013 (as the Australian Census Longitudinal Dataset, 2006-2011, (cat. no. 2080.0), bringing together a sample of almost one million records from the 2006 Census (Wave 1) with corresponding records from the 2011 Census (Wave 2). The addition of corresponding information from the 2016 Census (Wave 3) will expand our understanding of the dynamics and transitions that have been driving change in Australia since the 2006 Census.

As information from subsequent Censuses are added to the ACLD, its value as a resource for longitudinal studies of the Australian population will continue to increase.

AVAILABLE PRODUCTS

The following microdata products are available for this longitudinal dataset:

- ACLD 2011-2016 in TableBuilder - an online tool for creating tables and graphs.
- ACLD 2011-2016 in DataLab for in-depth analysis using a range of statistical software packages.

Further information about TableBuilder and DataLab, and other information to assist users in understanding and accessing microdata in general, is available from the Microdata Entry Page.

APPLYING FOR ACCESS

Before applying for access to TableBuilder, users should read and familiarise themselves with the information contained in the TableBuilder User Guide (cat. no. 1406.0.55.005).

To apply for access to TableBuilder, please register and apply from the How to Apply for Microdata page on the ABS website.

Information about TableBuilder can be found on the Microdata Entry Page.

If you already are a registered TableBuilder user Login in here.

To apply for access to the ACLD in DataLab, please contact Microdata Access Strategies via microdata.access@abs.gov.au.

Information about the DataLab can be found on the About the DataLab page.

SUPPORT

For support in the use of this product, please contact Microdata Access Strategies on 02 6252 7714 or via microdata.access@abs.gov.au.

DATA AVAILABLE ON REQUEST

Customised tables are available on a fee-for-service basis. A consultancy service is available for complex analysis and modelling. For further information, contact the National Information and Referral Service on 1300 135 070 or email client.services@abs.gov.au.

INQUIRIES

For further information about these and related statistics, contact the National Information and Referral Service on 1300 135 070, or email client.services@abs.gov.au. The ABS Privacy Policy outlines how the ABS will handle any personal information that you provide to us.

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016 Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

METHODOLOGY

SCOPE AND COVERAGE

The ACLD is a random 5% sample of persons enumerated in Australia on Census Night, 9 August 2011, which has been linked using statistical techniques to records of persons enumerated in Australia on Census Night, 9 August 2016.

The Census covers all areas in Australia and includes persons living in both private and non-private dwellings but excludes:

- diplomatic personnel of overseas governments and their families
- Australian residents overseas on Census Night and
- persons who expected to be usually resident in Australia for less than six months.

Overseas visitors are excluded from the 2011 ACLD sample. Visitors within Australia to private and non-private dwellings on Census Night are included.

For more information on the scope and coverage of the Census:

- How Australia Takes a Census, 2011 (cat. no. 2903.0)
- Census of Population and Housing: Understanding the Census and Census Data, Australia, 2016 (cat. no. 2900.0).

SAMPLE DESIGN

In preparation for adding 2016 Census data to the ACLD, a new panel of 2011 Census records was selected as a representative sample of the 2011 population. The 2011 Panel was designed to include most of the 2011 Census records that were linked in the 2006 Panel.

Sample maintenance

Without sample maintenance, the ACLD would decline in its ability to accurately reflect the Australian population over time, due to:

- people newly in scope of the ACLD (i.e. children born and immigrants who arrived in Australia since the previous Census) not being represented in the sample,
- people no longer being in scope due to death or overseas migration, and
- missing and/or incorrect links.

The 2011 Panel sample was increased slightly to 5.7%, to achieve a linked sample size of no greater than 5% of the population after allowing for missed links and people no longer being in scope due to death or overseas migration. The 2011 panel sample of over one million records (1,221,057) from the 2011 Census was linked to the 2016 Census, resulting in a linked sample size of 927,520 records at a linkage rate of 76%. This achieved a linked sample size of 4.3%.

Multi panel sample method

The ACLD sample is maintained through the application of the Multi-Panel framework, which provides an approach for selecting records in the ACLD to create panels that maintain the longitudinal and cross-sectional representativeness of the dataset over time, while minimising the impact of accumulated linkage bias on longitudinal analysis.

The Multi-Panel framework is comprised of multiple overlapping panels, with each panel representing a single Census population (2006, 2011, 2016, etc.). Each Census year panel will be linked to subsequent Censuses. The sample selection strategy for each panel is designed to maintain a linked sample size of 5%, maximise sample overlap between the panels, and introduce new records to the dataset in each panel to account for new births, migrants and missed links in previous panels.

This allows flexibility for users, who can draw on the most appropriate panel for their research question.

For further information on the Multi-Panel framework refer to Information Paper: Australian Census Longitudinal Dataset, Methodology and Quality Assessment, 2011-2016 (cat. no. 2080.5).

LINKING METHODOLOGY

Linking Strategy

Data from the 2011 ACLD Panel sample and the 2016 Census were brought together using data linkage techniques.

Data linkage is typically undertaken using deterministic and/or probabilistic methods. Both of these were used to create the 2011-2016 ACLD:

- Deterministic linkage: involves assigning record pairs across two datasets that match exactly or closely on common variables. This type of linkage is most applicable where the records from different sources consistently report sufficient information and can be an efficient process for conducting linkage.
- Probabilistic linkage: is based on the level of overall agreement on a set of variables common to the two datasets. This approach allows links to be assigned in spite of some missing or inconsistent information, providing there is enough agreement on other variables.

Linking Variables

The variables on the 2011 and 2016 Census files that were used for linking include:

- First name hash code
- Surname hash code
- Age

- Sex
- Date of birth
- Indigenous status
- Country of birth
- Year of arrival
- Marital status
- Religion
- Language spoken
- Mother's age
- Mother's day and month of birth
- Mother's country of birth
- Father's age
- Father's day and month of birth
- Father's country of birth
- Mesh block
- Statistical Areas 1, 2 and 4.

A number of linkage passes were conducted based on different combinations of these variables to ensure each record had the highest possible chance of being linked.

At the end of the linkage process, 927,520 (76%) of the 1,221,057 sample records from 2011 were linked to a 2016 Census record.

There were two reasons why some records from the 2011 Census were not linked to a record from the 2016 Census:

1. Records belonging to the same individual were present at both time points but these records failed to be linked because they contained missing or inconsistent information.
2. The person had no record in the 2016 Census.

For detailed information on the linking methodology and an assessment of its quality see Information Paper: Australian Census Longitudinal Dataset, Methodology and Quality Assessment, 2011-2016 (cat. no. 2080.5).

To protect the privacy of Census respondents, we used an ABS encoded Census name for linking 2011 and 2016 Census records in the ACLD. Encoding was undertaken in 2011 for the purpose of protecting privacy by anonymising name and improving the future quality and efficiency of the linking process.

The codes are created by grouping people with a combination of letters from their first and last names using a secure one-way process, meaning that a code cannot be reversed to deduce the original name information. Each code represents approximately 2,000 people drawn from many different letter combinations, and therefore is not unique to an individual. Actual name information from the 2016 Census was not used to link to 2011 Census records.

For further information, see Information Paper: Australian Census Longitudinal Dataset, Methodology and Quality Assessment, 2011-2016 (cat. no. 2080.5).

WEIGHTING, BENCHMARKING AND ESTIMATION

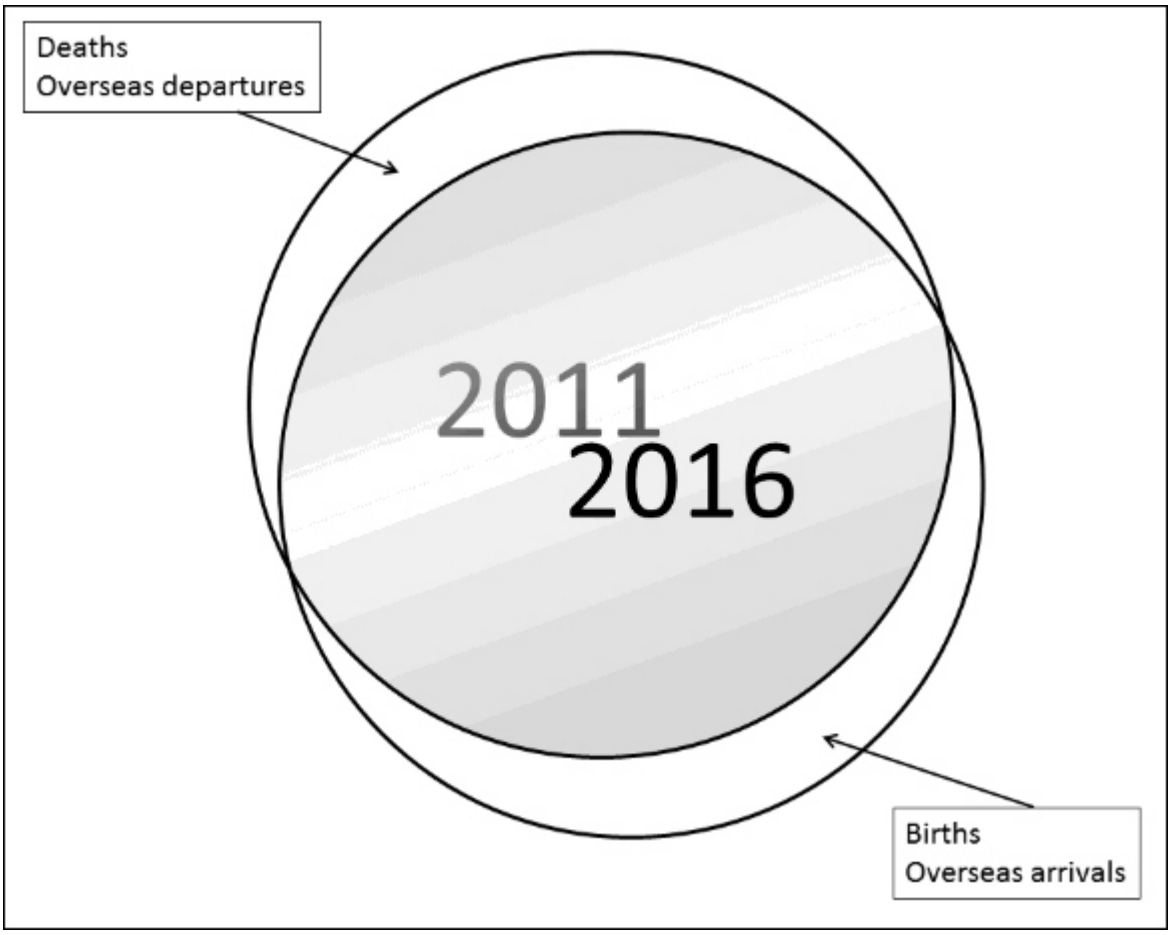
Weighting

Weighting is the process of adjusting a sample to infer results for the relevant population. To do this, a 'weight' is allocated to each sample unit - in this case, persons. The weight can be considered an indication of how many people in the relevant population are represented by each person in the sample. Weights were created for linked records in the ACLD to enable longitudinal population estimates to be produced.

The 2011 Panel of the ACLD is a random 5% sample of persons enumerated in Australia on Census Night, 9 August 2011. As such, each person in the sample should represent about 20 people in the Australian population. Between Censuses, however, the Australian population in scope of the ACLD changes as people die or move overseas. In addition, Census net undercount and data quality can affect the capacity to link equivalent records across waves.

The ACLD weights benchmark the linked records to the estimated Australian population that was in scope of both the 2011 and 2016 Censuses. The weights were based on four components: the design weight, undercoverage adjustment, missed link adjustment and population benchmarking.

The population benchmark is based on the 2016 Estimated Resident Population (ERP). This population benchmark was adjusted by the estimated probability that a person was also in Australia in 2011. This probability was estimated using the 2016 Census 'reported 5 year ago address' variable.



Weights were benchmarked to the following population groups:

- state/territory by age (ten year groups) by sex by mobility (interstate arrivals benchmarked separately), and
- Indigenous status by state/territory.

The mean final weight for the linked records is 22.3 for females and 23.2 for males. The weights

range between 14.8 and 83. The mean weight was higher for Aboriginal and Torres Strait Islander persons and for people in the Northern Territory.

Estimation

Estimates of population groups are obtained by summing the weights of persons with the characteristic(s) of interest.

For further information about ACLD weighting and estimation refer to the Information Paper: Australian Census Longitudinal Dataset, Methodology and Quality Assessment, 2011-2016 (cat. no. 2080.5).

SOURCES OF ERROR

All reasonable attempts have been taken to ensure the accuracy of the longitudinal dataset. Nevertheless potential sources of error including sampling, linking and Census quality error should be kept in mind when interpreting the results.

Sampling Error

Sampling error occurs because only a small proportion of the total population is used to produce estimates that represent the whole population. Sampling error refers to the fact that for a given sample size, each sample will produce different results, which will usually not be equal to the population value.

There are two common ways of reducing sampling error - increasing sample size and/or utilising an appropriate selection method (for example, multi-stage sampling would be appropriate for household surveys). Given the large sample size for the ACLD (1 in 20 persons), and simple random selection, sampling error is minimal.

Linking Accuracy

False links can occur during the linkage process as even when a record pair matches on all or most linking fields, it may not actually belong to the same individual. While the methodology is designed to ensure that the vast majority of links are true, some are nevertheless false. The nature of the process used for the ACLD linkage means that while the links obtained are to a high degree of accuracy, some false links may be present within the ACLD dataset. There is an estimated 1% false link rate in the ACLD.

For further detail on the accuracy of the linkage, see Information Paper: Australian Census Longitudinal Dataset, Methodology and Quality Assessment, 2011-2016 (cat. no. 2080.5).

Managing Census Quality

The ABS aims to produce high quality data from the Census. To achieve this, extensive effort is put into Census form design, collection procedures, and processing procedures.

There are four principle sources of error in Census data: respondent error, processing error, partial response and undercount. Quality management of the Census program aims to reduce error as much as possible, and to provide a measure of the remaining error to data users, to allow them to use the data in an informed way.

Information about the quality of the 2011 Census data is available on the Data Quality page on the ABS website.

The Census Independent Assurance Panel concluded that the 2016 Census data is of comparable quality to 2011 and 2006 Census data, so may be used with confidence. Information is available in Census of Population and Housing: Understanding the Census and Census Data, Australia, 2016 (cat. no. 2900.0).

For more detail see Managing Census Quality, in Census of Population and Housing: Census Dictionary, 2016 (cat. no. 2901.0).

Respondent Error

For most households in Australia, the Census is self-enumerated. This means that householders are required to complete the Census form themselves, rather than having the help of a Census collector. The Census form may be completed by one household member on behalf of others. Error can be introduced if the respondent does not understand the question, or does not know the correct information about other household members. Self-enumeration carries the risk that wrong answers could be given, either intentionally or unintentionally

Processing Error

Much of the data on the Census form is recorded using automatic processes, such as scanning, Intelligent Character Recognition and other automatic processes. Quality assurance procedures are used during Census processing to ensure processing errors are kept at an acceptable level. Sample checking is undertaken during coding operations, and corrections are made where necessary.

Partial Response

When completing their Census form, some people do not answer all the questions which apply to them. While questions of a sensitive nature are generally excluded from the Census, all topics have a level of non-response. This can be measured and is generally low. In those instances where a householder fails to answer a question, a 'not stated' code is allocated during processing, with the exception of non-response to age, sex, marital status and place of usual residence. These variables are needed for population estimates, so they are imputed using other information on the Census form, as well as information from the previous Census.

Undercount

The goal of the Census is to obtain a complete measure of the number and characteristics of people in Australia on Census Night and their dwellings, but it is inevitable that a small number will be missed and some will be counted more than once. In Australia more people are missed from the Census than are counted more than once, thus the effect when both factors are taken into account is a net undercount.

For more detail see Managing Census Quality, in Census of Population and Housing: Census Dictionary, 2016 (cat. no. 2901.0).

DATA CONSISTENCY

A small percentage of linked records have inconsistent data, such as a different country of birth at the two time points or an age inconsistency of more than one year (when the expected five year difference is accounted for). Inconsistencies may be due to:

- false link - the record pair does not belong to the same individual

- reporting error - information for the same individual was reported differently in 2011 and in 2016
- processing error - the value of a data item was inaccurately assigned or imputed during processing.

In most analysis, the effect of inconsistent information may only have a small impact. Characteristics from the 2011 or the 2016 data can be used in tables and some exploration of consistency over time will assist in drawing appropriate conclusions.

No data editing was applied to the file beyond that which had already taken place during the relevant Census processing period. A set of consistency flags has been included on the ACLD file so that inconsistent data may be observed, quantified or excluded from calculations. Consistency flags, located in the Quality Indicators group of data items, have been created for Census variables that would not be expected to change over time or have unlikely transitions over time. These are as follows:

- Age
- Sex
- Birthplace of Person
- Birthplace of Female Parent
- Birthplace of Male Parent
- Year of Arrival
- Indigenous Status
- Registered Marital Status
- Highest Year of School Completed
- Level of Highest Non-School Qualification
- Country of Birth of Spouse or Partner
- Number of Children Ever Born.

There are numerous ways to define 'consistency'. The consistency flags have fine level categories to allow users flexibility in using their own definition of 'consistent' or 'inconsistent'. For example, where one Census has 'not stated' for the year of arrival data item, a user can decide whether the record should be considered consistent or not. The same applies to where the response for one Census is 'not applicable'. The labels attached to each category suggesting consistency or inconsistency will assist the user in determining which records are consistent or inconsistent for their needs.

See also Quality Indicators in the Data Items section.

INCONSISTENT REPORTING ON THE LINKED ACLD FILE, By Selected Characteristics

Characteristic	Proportion of linked records with inconsistent data between 2011 and 2016
Age (within 1 year)	0.9%
Sex	0.2%
Birthplace of person	1.0%
Birthplace of female parent	1.5%
Birthplace of male parent	1.9%
Year of arrival	4.0%
Indigenous status	2.2%
Registered marital status	0.4%
Highest year of school completed	5.7%
Level of highest non-school qualification	6.9%

Country of birth of spouse or partner	0.6%
Number of children ever born	0.5%

This page last updated 19 March 2019

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016

Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

FILE STRUCTURE

The ACLD datasets in both TableBuilder and DataLab are single-level files that count persons.

In TableBuilder, data items have first been separated by year of collection. Each data item for the person in 2011 has a corresponding item for 2016. In addition, data items that provide longitudinal information derived from both years, such as the consistency in reporting of certain data items between 2011 and 2016, is in a folder labelled 'Quality Indicators' (see also Using the ACLD in TableBuilder).

For each year, data items are further divided into the following groups:

- **Person** - person-level characteristics, covering demographic, culture and language, education, employment and unpaid work topics.
- **Family** - derived from the characteristics of the family to which the person belongs on Census Night - e.g., age of youngest child in family.
- **Household** - derived from the characteristics of the household to which the person belongs on Census Night - e.g., total weekly household income.
- **Dwelling** - derived from the characteristics of the dwelling in which the person was enumerated - e.g., type of dwelling.
- **Geographical Area** - geographic classifications on both a usual residence and place of enumeration basis.
- **Female Parent** - derived from information about the person's female parent, provided the parent was counted in the same family on Census Night - e.g., labour force status of female parent.
- **Male Parent** - derived from information about the person's male parent, provided the parent was counted in the same family on Census Night - e.g., labour force status of male parent.
- **Spouse/Partner** - derived from information about the person's spouse/partner, provided the spouse/partner was counted in the same family on Census Night - e.g., highest educational attainment of spouse/partner.
- **Quality Indicators** - consistency flags, imputation flags and movement indicators to assist with the assessment of the suitability of records for inclusion in analysis.

UNLINKED RECORDS

The ACLD file contains all sample records from the 2011 Census panel, both those that were linked to a 2016 record and those that were not. Records that were linked have information for 2011 and 2016. Records that were not linked have information for 2011 only.

Data items for both 2011 and 2016 have a category labelled: "Unlinked record".

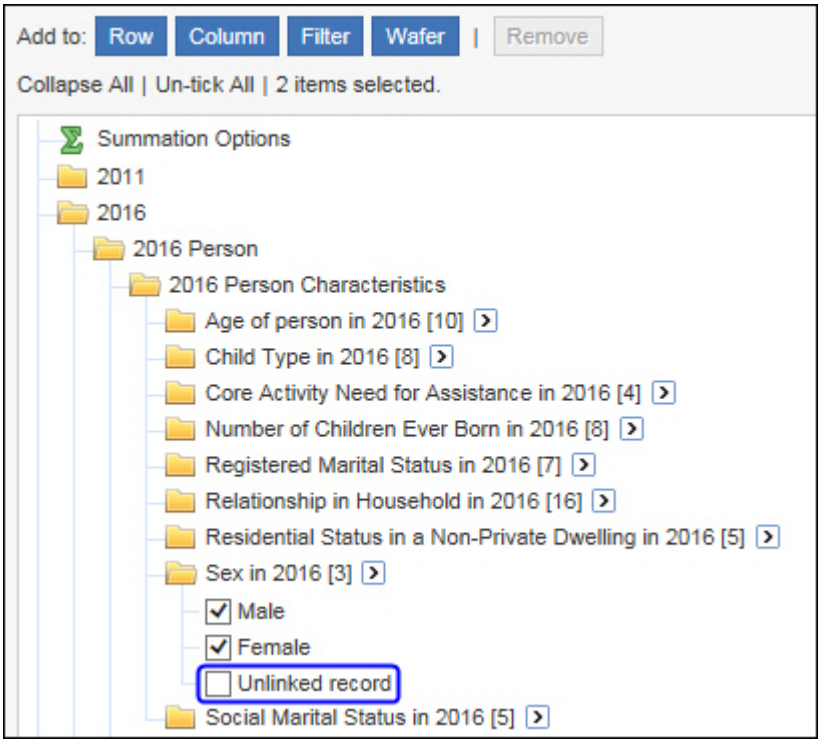
- the unlinked record category for 2011 data items is empty. It may take on a value in the

future once the file is further augmented.
- the unlinked record category for 2016 data items indicates records from the 2011 Census that were not linked to a corresponding 2016 Census record, i.e., these records have valid information for 2011 but no information for 2016.

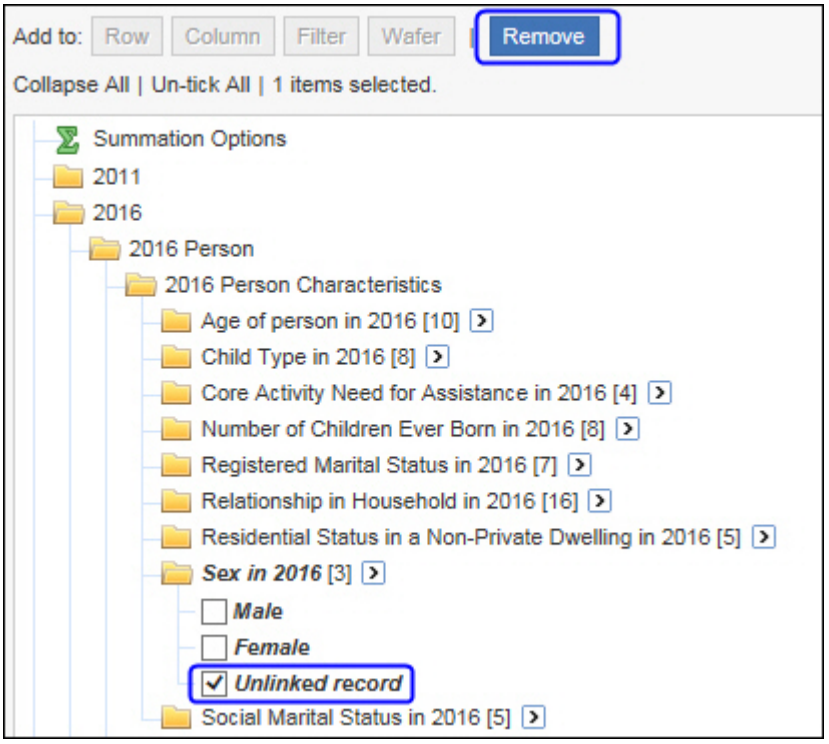
Excluding Unlinked Records in TableBuilder

When using the weighted summation option in TableBuilder, no results will be returned for unlinked records in 2016, as weights were not applied to these records. Results including unlinked 2011 records will only be returned if analysis is performed on unweighted data.

To exclude unlinked records from your analysis, deselect the "Unlinked record" category in each data item before adding it to the table. Such a table would produce a sample count corresponding to the equivalent table run with weights. Refer to the TableBuilder User Guide (cat. no. 1406.0.55.005) for more information on how to select data items for tables.



If the 'unlinked record' category is present on a data item that has already been added to a table, it can be removed by selecting this category within the relevant data item and then pressing the 'Remove from Table' button.



Note that removing any category, such as the 'Unlinked record' category, from a table where data has already been generated will clear all data, meaning the table will need to be rerun.

This page last updated 19 March 2019

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016 Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

USING THE ACLD IN TABLEBUILDER

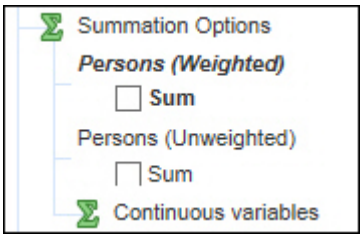
TABLEBUILDER USER GUIDE

The TableBuilder User Guide (cat. no. 1406.0.55.005) is a comprehensive reference guide for the web interface of TableBuilder. It includes information on building and working with tables, customising data, understanding the results, data visualisation options, and confidentiality processes.

COUNTING UNITS AND WEIGHTS

Weighting is the process of adjusting results from a sample to infer results for the total population. To do this, a weight is allocated to each person. The weight is the value that indicates how many population units are represented by the sample unit.

Both the sample and weighted count options have been made available for the ACLD. It is therefore critical that weighted or unweighted counts are selected as appropriate when specifying tables. The following image shows the available Summation Options.



The default option used for the ACLD is weighted count. Weights should be used when making inferences about the longitudinal Australian population and will be the basis for most analyses. Uses for unweighted counts are generally limited to research into unlinked records and more sophisticated analysis for those seeking to understand the weighting methodology better or wishing to apply their own weighting methods.

RELATIVE STANDARD ERROR

While weighted counts are available in the ACLD TableBuilder, the Relative Standard Error will not be calculated for these counts due to the confounding effects of linking error present in the sample, which were not able to be quantified.

CONFIDENTIALITY FEATURES IN TABLEBUILDER

In accordance with the Census and Statistics Act 1905, all the data in TableBuilder are subjected to a confidentiality process before release. This confidentiality process is undertaken to avoid releasing information that may allow the identification of particular individuals, families, households, dwellings or businesses.

Processes used in the 2011-2016 ACLD in TableBuilder to confidentialise records include the following:

- perturbation of data
- table suppression.

Perturbation of data

To minimise the risk of identifying individuals in aggregate statistics, a technique is used to randomly adjust cell values. This technique is called perturbation. Perturbation involves small random adjustments of the statistics and is considered the most satisfactory technique for avoiding the release of identifiable statistics while maximising the range of information that can be released. These adjustments have a negligible impact on the underlying pattern of the statistics.

The introduction of these random adjustments result in tables not adding up. Randomly adjusted individual cells will be consistent across tables, but the totals in any table may not be the sum of the individual cell values. The size of the difference between summed cells and the relevant total will generally be very small, as demonstrated below.

Persons (Unweighted)	Persons (Unweighted) #
Sex in 2016	
Male	450,054.0
Female	477,460.0
Total	927,520.0

(Sum of cells = 450,054 + 477,460 = 927,514. Difference of 6 relative to displayed total.)

Table suppression

Some tables generated within TableBuilder may contain a substantial proportion of very low counts within cells (excluding cells that have counts of zero). When this occurs, all values within the table are suppressed in order to preserve confidentiality. The following error message displayed at the bottom of the table indicates when table suppression has occurred.

'ERROR: The table has been suppressed as it is too sparse'.

ACCESS TO TABLEBUILDER

To access the ACLD via TableBuilder, please register or log in, via the Microdata Entry Page. Please familiarise yourself with the Responsible Use of ABS Microdata Guide (cat. no. 1406.0.55.003), if you intend to access ACLD microdata.

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016

Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

THE ACLD IN THE DATALAB

The DataLab is an interactive data analysis solution available for high end users to run advanced multivariate statistical analyses, for example, multiple regressions and structural equation modelling. The DataLab environment contains up-to-date versions of SPSS, Stata, SAS and R analytical languages. Controls in the DataLab have been put in place to protect the identification of individuals and organisations. These controls include environmental protections, data de-identification and confidentialisation, access safe guards and output clearance. All output from DataLab sessions is cleared by an ABS officer before it is released.

For more information about the DataLab please see About the DataLab. To apply for access to the ACLD in the DataLab, please contact Microdata Access Strategies via microdata.access@abs.gov.au.

For information about accessing the ACLD test file for DataLab, please see ABS DataLab ACLD Test File.

To access the ACLD via the ABS DataLab, please register or log in, via the Microdata Entry Page. Please familiarise yourself with the Responsible Use of ABS Microdata Guide (cat. no. 1406.0.55.003), if you intend to access ACLD microdata.

This page last updated 19 March 2019

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016

Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

ABS DATALAB ACLD TEST FILE

Note: The test file does not contain real data and cannot be used for analysis.

A test file has been created for the ACLD microdata product, which is consistent with the structure outlined in File Structure. The purpose of the test file is to allow researchers/analysts to become familiar with the data structure and prepare code/programs prior to applying for, or commencing, a DataLab session. This aims to maximise the value of sessions by saving users' time and resources once they enter the DataLab environment.

The test file mimics the structure of the ACLD microdata datasets - it has the same data items and allowed values, however, all data in the test file is false, created through a randomisation process. Proportions of values within data items in the test file will be similar to those in the real data; however, relationships between data items are not (intentionally) maintained. It is extremely unlikely that a record in the test file would match with a genuine record in the real data.

The test file is available as a free download from the Downloads tab. It can also be made available in other file formats on request, if required. For further information users should email microdata.access@abs.gov.au.

This page last updated 19 March 2019

DATA ITEMS

DATA ITEMS LIST

A complete list of data items included on the ACLD file is provided in an Excel spreadsheet that can be accessed from the Downloads tab.

All data items are created at the person level. This includes data items relating to the family and household of the person selected in the sample. For ease of use, these data items have been divided into Person, Dwelling, Household, Family, Spouse Related, and Male and Female parent related groupings.

Users intending to subscribe to the TableBuilder product or use the microdata product in the DataLab should ensure the data they require, and the level of detail required, are available and applicable for the intended use.

For information about the quality of the data items in the ACLD please see the Quality Declaration.

Note: While the 2011 and 2016 Censuses were predominantly the same, there were some minor differences. For example, a number of changes were made to how industry of employment information was collected for the 2016 Census. The ABS advises this data is not directly comparable to 2011 industry data and should not be used to measure longitudinal transitions. For further information refer to Industry of Employment (INDP) in Census of Population and Housing: Understanding the Census and Census Data, Australia, 2016 (cat. no. 2900.0).

NEW DATA ITEMS

The new data items for the 2011-16 ACLD are as follows:

NEW DATA ITEMS

Mnemonic	Data Item
Both TableBuilder and DataLab	
BPFPR	Country of Birth of Mother
BPMPR	Country of Birth of Father
CPUNCHP	Number of Persons in Household who Provide Assistance to a Person with Disability as stated
EETP	Engagement in Employment, Education and Training
ENGLP	Proficiency in Spoken English/Language
FBLFP	Family Blending
IFAGEP	Imputation Flag for Age
IFMSTP	Imputation Flag for Registered Marital Status
IFPURP	Imputation Flag for Usual Address
IFSEXP	Imputation Flag for Sex
INGDWTDP	Indigenous Dwelling Indicator
LFSFP	Labour Force Status of Parents/Partners in Families
SIEMP	Status in Employment
TENLLDP	Tenure and Landlord Type
DataLab Only	
MBUCD	Mesh Block of Enumeration
MBUCP	Mesh Block of Usual Residence
SA1UCD	Statistical Area Level 1 of Enumeration
SA1UCP	Statistical Area Level 1 of Usual Residence

VISITORS ON CENSUS NIGHT

Overseas visitors were excluded from the 2011 ACLD sample. The ACLD, however, does include visitors from within Australia. These are people who were enumerated away from their usual residence on Census Night. Family information cannot be derived for these persons and as such, all family, spouse, and male and female parent related data items are not applicable for visitors.

All dwelling related data items, however, have been made applicable to visitors. This information relates to their dwelling of

enumeration on Census Night, not usual residence.

Most household data items are not applicable to visitors, however for four data items, visitors have been included in order to align to standard Census derivations of that data item. These comprise:

- Total Household Income as stated (weekly) of household in which person was enumerated
- Total Household Income (weekly) of household in which person was enumerated
- Household Income Derivation Indicator of household in which person was enumerated
- Household Composition of household in which person was enumerated.

Any applicable household information for a visitor relates to their place of enumeration, not usual residence.

Where a data item is also applicable to visitors, the usual address indicator data item for the relevant Census year can be used to restrict the table to usual residents only.

The cell comments available in the data item list provide precise information on who is, and is not, applicable for each data item.

PERSONS TEMPORARILY ABSENT ON CENSUS NIGHT

The Census household form provides the opportunity to list up to three persons who were temporarily absent from the dwelling on Census Night. A limited amount of information is collected for these persons and it is used to better derive the family and household characteristics of the dwelling. In deriving family and household related data items for the ACLD, information on persons temporarily absent was included where relevant and available. Details are provided in cell comments in the data items list.

NOT APPLICABLE CATEGORIES

Most data items in the ACLD include a 'not applicable' category. The definition of the 'not applicable' category, where relevant, can be found in the section of the relevant year's Census Dictionary (cat. no. 2901.0), including the previous year releases found on the "Past & Future Releases" tab.

NOT STATED CATEGORIES

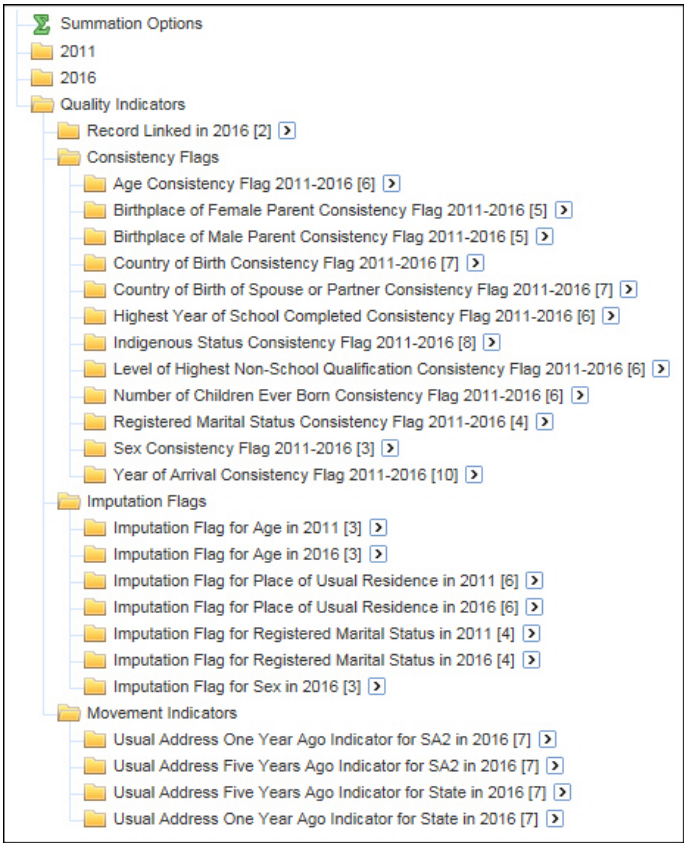
'Not stated' categories occur when no response has been provided for a data item. All Census data items contain 'not stated' categories except for age, sex, marital status and usual address, as this information is imputed for these items.

Other Census products commonly use the symbol '&' to denote a code value of 'not stated'. In the ACLD, the symbol 'X' or '97' has been used. The codes are listed in the data items list.

QUALITY INDICATORS

The ACLD contains a number of data items that relate to the quality of linkage over the period 2011 to 2016 and have been collectively named Quality Indicators. The first of these are consistency flags. These variables measure the consistency of reporting on linked records between 2011 and 2016. The following consistency flags can be found in the Quality Indicators folder of the TableBuilder data item tree and in the ACLD microdata product available in the DataLab:

- Age
- Sex
- Birthplace of Person
- Birthplace of Female Parent
- Birthplace of Male Parent
- Year of Arrival
- Indigenous Status
- Registered Marital Status
- Highest Year of School Completed
- Level of Highest Non-School Qualification
- Country of Birth of Spouse or Partner
- Number of Children Ever Born.



Consistency flags can be used with other variables. For example, age inconsistency can be cross tabulated with sex to examine potential gender differences in the reporting of age.

Persons (Weighted)	Persons (Weighted) #		
Sex in 2016	Male	Female	Total
Age Consistency Flag 2011-2016			
Unlinked record	0.0	0.0	0.0
Age consistent within 0 years	10,052,131.4	10,259,414.9	20,311,545.9
Age inconsistent within 1 year	292,365.4	281,620.3	573,960.2
Age inconsistent from 2 to less than 5 years	38,352.7	37,589.0	75,874.1
Age inconsistent from 5 to less than 10 years	25,271.9	27,378.7	52,647.7
Age inconsistent by 10 or more years	32,384.8	33,826.1	66,237.6
Total	10,440,364.4	10,639,712.6	21,080,213.8

In addition to the consistency flags, a "Record linked in 2016 Flag" is also available in the Quality Indicators folder. This flag can be cross tabulated with another data item to examine linkage rates (that is, the proportion of records linked). For example, cross tabulating the record linked flag with State/Territory of usual residence enables an examination of differences in linkage rates between the states and territories.

Record Linked in 2016	2011 to 2016 record linked	2011 record not linked to 2016	Total
<< Australian Statistical Geography Standard (Usual Residence) (Statistical Areas) in 2011			
New South Wales	75.93%	24.07%	100.00%
Victoria	76.72%	23.28%	100.00%
Queensland	74.87%	25.13%	100.00%
South Australia	78.26%	21.74%	100.00%
Western Australia	75.77%	24.23%	100.00%
Tasmania	76.39%	23.64%	100.00%
Northern Territory	62.26%	37.75%	100.00%
Australian Capital Territory	76.61%	23.36%	100.00%
Other Territories	53.93%	47.75%	100.00%
Total	75.96%	24.04%	100.00%

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016

Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

ANALYSING THE ACLD - UNDERSTANDING CONSISTENCY BETWEEN WAVES

The ACLD is a longitudinal dataset using data from successive Censuses.

While the 2011 and 2016 Censuses had predominantly the same questions and were processed in a similar way, there were some differences between them.

Users are encouraged to read Understanding the data pages in Census of Population and Housing: Understanding the Census and Census Data, Australia, 2016 (cat. no. 2900.0), and information in How Australia Takes a Census, 2011 (cat. no. 2903.0) for information to assist with using and interpreting specific data items across time.

For example, a number of changes were made to how industry of employment information was collected for the 2016 Census. The ABS advises this data is not directly comparable to the previous Census Industry of employment data, and should not be used to measure longitudinal transitions between industries from 2011 to 2016. For further information refer to Industry of Employment (INDP) in Census of Population and Housing: Understanding the Census and Census Data, Australia, 2016 (cat. no. 2900.0).

Other data items that are different between Census years are personal, family and household income. Income was collected in ranges and these ranges are different in different Census years. The ACLD does not include an adjustment to income data for inflation.

A small percentage of linked records have inconsistent data, such as a different country of birth at the two time points or an age inconsistency of more than one year. Inconsistencies may be due to:

- false link - the record pair does not belong to the same individual
- reporting error - information for the same individual was reported differently in 2011 and in 2016
- processing error - the value of a data item was inaccurately assigned or imputed during processing.

For further information about the ACLD linkage methodology refer to Information Paper: Australian Census Longitudinal Dataset, Methodology and Quality Assessment, 2011-2016 (cat. no. 2080.5).

ACLD microdata contains a large number of data items and in some cases the level of detail has been collapsed from that described in the Census Dictionary. For more information on the level of detail provided, please see the associated Data Items list.

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016

Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

The Australian Census Longitudinal Dataset (ACLD) uses data from the Census of Population and Housing to build a rich longitudinal picture of Australian society. The ACLD can uncover new insights into the dynamics and transitions that drive social and economic change over time, and how these vary for diverse population groups and geographies.

The ACLD is a random five per cent sample of the Australian population and three waves of data have so far contributed to the ACLD from the 2006 Census, 2011 Census and 2016 Census.

In this release of the 2011-2016 ACLD, a representative sample of over 1.2 million records from the 2011 Census was brought together with corresponding records from the 2016 Census to form the 2011 Panel of the ACLD. The 2011 Panel includes new births and migrants since the 2006 Census, and is a rich source for exploring how Australian society has changed between the 2011 and 2016 Censuses.

The information within the Summary tab provides an overview of the methodology, file structure, data items and how to use the microdata files related to this release. For more detailed information about the sampling and linking methodology and results, refer to Information Paper: Australian Census Longitudinal Dataset, Methodology and Quality Assessment, 2011-2016 (cat. no. 2080.5).

A detailed list of data items is available on the Downloads tab.

The following microdata products are available:
- 2011-2016 ACLD in TableBuilder
- 2011-2016 ACLD in the DataLab

For information about:
- TableBuilder, refer to: https://www.abs.gov.au/websitedbs/D3310114.nsf/home/About+TableBuilder
- DataLab, refer to: https://www.abs.gov.au/websitedbs/D3310114.nsf/home/CURF:+About+the+ABS+Data+Laboratory+(ABSDL)

For access to the ACLD, apply online at: www.abs.gov.au/registration

Approved users can use the microdata to create their own customised tables using weighted and unweighted data.

2080.0 - Microdata: Australian Census Longitudinal Dataset, 2011-2016

Quality Declaration

ARCHIVED ISSUE Released at 11:30 AM (CANBERRA TIME) 27/02/2018

This document was added or updated on 06/03/2018.

24/10/2018 - TableBuilder updated:

- 'Equivalised Total Household Income (weekly) of household in which person was enumerated in 2016' corrected and replaced
- 'Age Consistency Flag 2011-2016' corrected and replaced
- 'Indigenous Status Consistency Flag 2011-2016' corrected and replaced
- Continuous data items added

30/08/2018 - DataLab file updated:

- 'Age Consistency Flag 2011-2016' corrected and replaced
- 'Indigenous Status Consistency Flag 2011-2016' corrected and replaced

29/05/2018 - DataLab file updated:

- 'Equivalised Total Household Income (weekly) of household in which person was enumerated in 2016' corrected and replaced

29/03/2018 - Minor formatting improvements to Data Item Lists.

06/03/2018 - Minor formatting improvements and fixes to links. No changes to data.

This page last updated 19 March 2019